

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/612,314

DATE: 11/14/2001

TIME: 09:58:50

Input Set : A:\379450004.app

Output Set: N:\CRF3\11142001\I612314.raw

3 <110> APPLICANT: SMITH, Richard, Anthony, Godwin  
 4 DODD, Ian  
 5 MOSSAKOWSKA, Danuta, Ewa, Irena  
 7 <120> TITLE OF INVENTION: CONJUGATES OF SOLUBLE PEPTIDIC COMPOUNDS WITH  
 8 MEMBRANE-BINDING AGENTS  
 10 <130> FILE REFERENCE: 37945-0004  
 12 <140> CURRENT APPLICATION NUMBER: US 09/612,314  
 13 <141> CURRENT FILING DATE: 2000-07-07  
 15 <150> PRIOR APPLICATION NUMBER: US 09/214,913  
 16 <151> PRIOR FILING DATE: 1999-03-16  
 18 <150> PRIOR APPLICATION NUMBER: PCT/EP97/03715  
 19 <151> PRIOR FILING DATE: 1997-07-08  
 21 <150> PRIOR APPLICATION NUMBER: GB 96 148 71.3  
 22 <151> PRIOR FILING DATE: 1996-07-15  
 24 <160> NUMBER OF SEQ ID NOS: 67  
 26 <170> SOFTWARE: PatentIn Ver. 2.1  
 28 <210> SEQ ID NO: 1  
 29 <211> LENGTH: 37  
 30 <212> TYPE: DNA  
 31 <213> ORGANISM: Artificial Sequence  
 32 <220> FEATURE:  
 34 <223> OTHER INFORMATION: Description of Artificial Sequence:  
 35 Oligonucleotide used to anneal to oligonucleotide  
 36 of SEQ ID NO.2  
 38 <400> SEQUENCE: 1  
 39 gcaccgcagt gcatacatccc gaacaaatgc taataaa 37  
 41 <210> SEQ ID NO: 2  
 42 <211> LENGTH: 37  
 43 <212> TYPE: DNA  
 44 <213> ORGANISM: Artificial Sequence  
 45 <220> FEATURE:  
 47 <223> OTHER INFORMATION: Description of Artificial Sequence:  
 48 Oligonucleotide used to anneal to oligonucleotide  
 49 of SEQ ID NO.1  
 51 <400> SEQUENCE: 2  
 52 agctttttatt agcatttgtt cgggatgatg caactgcg 37  
 54 <210> SEQ ID NO: 3  
 55 <211> LENGTH: 85  
 56 <212> TYPE: DNA  
 57 <213> ORGANISM: Artificial Sequence  
 58 <220> FEATURE:  
 60 <223> OTHER INFORMATION: Description of Artificial Sequence:  
 61 Oligonucleotide used to anneal to oligonucleotide  
 62 of SEQ ID NO.4 to generate fragment 4  
 64 <400> SEQUENCE: 3  
 65 gcaccgcagt gcatacatccc gaacaaagac ggtccgaaaa agaagaaaaa gaaatctccg 60  
 66 tccaaatctt ccggttgcta ataaa 85

ENTERED

RECEIVED

NOV 23 2001

TECH CENTER 1600/2900

## RAW SEQUENCE LISTING

DATE: 11/14/2001

PATENT APPLICATION: US/09/612,314

TIME: 09:58:50

Input Set : A:\379450004.app

Output Set: N:\CRF3\11142001\I612314.raw

RECEIVED

NOV 23 2001

TECH CENTER 1600/2900

```

68 <210> SEQ ID NO: 4
69 <211> LENGTH: 85
70 <212> TYPE: DNA
71 <213> ORGANISM: Artificial Sequence
73 <220> FEATURE:
74 <223> OTHER INFORMATION: Description of Artificial Sequence:
75     Oligonucleotide used to anneal to oligonucleotide
76     of SEQ ID NO.3 to generate fragment 4
78 <400> SEQUENCE: 4
79 agcttttatt agcaaccgga agatttggac ggagatttct ttttcttctt tttcggaccg 60
80 tctttgttcg ggaatgacga ctgcg                                     85
82 <210> SEQ ID NO: 5
83 <211> LENGTH: 17
84 <212> TYPE: PRT
85 <213> ORGANISM: Artificial Sequence
87 <220> FEATURE:
88 <221> NAME/KEY: UNSURE
89 <222> LOCATION: (17)
90 <223> OTHER INFORMATION: NH2 group is linked to the C-terminal cysteine
92 <220> FEATURE:
93 <223> OTHER INFORMATION: Description of Artificial Sequence: Peptide used
94     to synthesize MSWP-1
96 <400> SEQUENCE: 5
97 Gly Ser Ser Lys Ser Pro Ser Lys Lys Lys Lys Lys Pro Gly Asp
98   1           5           10           15
100 Cys
103 <210> SEQ ID NO: 6
104 <211> LENGTH: 198
105 <212> TYPE: PRT
106 <213> ORGANISM: Artificial Sequence
108 <220> FEATURE:
109 <223> OTHER INFORMATION: Description of Artificial Sequence: The petide
110     sequence correpsonds to short consensus repeats
111     1-3 of CRI with a C-terminal cysteine
113 <400> SEQUENCE: 6
114 Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
115   1           5           10           15
117 Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
118           20           25           30
120 Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
121           35           40           45
123 Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
124           50           55           60
126 Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
127   65           70           75           80
129 Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
130           85           90           95
132 Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
133           100           105           110

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/612,314

DATE: 11/14/2001  
 TIME: 09:58:50

Input Set : A:\379450004.app  
 Output Set: N:\CRF3\11142001\I612314.raw

```

135 Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
136      115      120      125
138 Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
139      130      135      140
141 Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
142 145      150      155      160
144 Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
145      165      170      175
147 Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
148      180      185      190
150 Ile Ile Pro Asn Lys Cys
151      195
154 <210> SEQ ID NO: 7
155 <211> LENGTH: 214
156 <212> TYPE: PRT
157 <213> ORGANISM: Artificial Sequence
159 <220> FEATURE:
160 <223> OTHER INFORMATION: Description of Artificial Sequence: Short
161 consensus repeats 1-3 of CRI with an additional 17
162 C-terminal amino acids
164 <400> SEQUENCE: 7
165 Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
166 1      5      10      15
168 Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
169      20      25      30
171 Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
172      35      40      45
174 Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
175      50      55      60
177 Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
178 65      70      75      80
180 Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
181      85      90      95
183 Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
184      100      105      110
186 Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
187      115      120      125
189 Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
190      130      135      140
192 Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
193 145      150      155      160
195 Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr
196      165      170      175
198 Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
199      180      185      190
201 Ile Ile Pro Asn Lys Asp Gly Pro Lys Lys Lys Lys Lys Lys Ser Pro
202      195      200      205
204 Ser Lys Ser Ser Gly Cys
205      210

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/612,314

DATE: 11/14/2001

TIME: 09:58:50

Input Set : A:\379450004.app

Output Set : N:\CRF3\11142001\I612314.raw

```

208 <210> SEQ ID NO: 8
209 <211> LENGTH: 215
210 <212> TYPE: PRT
211 <213> ORGANISM: Artificial Sequence
213 <220> FEATURE:
214 <221> NAME/KEY: UNSURE
215 <222> LOCATION: (198)..(199)
216 <223> OTHER INFORMATION: Residues 1-198 are a first polypeptide chain and
217 residues 199-215 are a second polypeptide chain
218 linked by a disulphide bond formed between the
219 cysteines at positions 198 and 199
221 <220> FEATURE:
222 <221> NAME/KEY: UNSURE
223 <222> LOCATION: (215)
224 <223> OTHER INFORMATION: The C-terminal glycine is linked to an
225 N-(Myristoyl) group
227 <220> FEATURE:
228 <221> NAME/KEY: UNSURE
229 <222> LOCATION: (198)
230 <223> OTHER INFORMATION: The cysteine at position 198 is -Cys-COOH
232 <220> FEATURE:
233 <221> NAME/KEY: UNSURE
234 <222> LOCATION: (199)
235 <223> OTHER INFORMATION: The cysteine at position 199 is CONH2-Cys-
237 <220> FEATURE:
238 <223> OTHER INFORMATION: Description of Artificial Sequence:
239 [SCR1-3]-Cys-S-S-[MSWP-1]
241 <400> SEQUENCE: 8
242 Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
243 1 5 10 15
245 Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
246 20 25 30
248 Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
249 35 40 45
251 Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
252 50 55 60
254 Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
255 65 70 75 80
257 Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
258 85 90 95
260 Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
261 100 105 110
263 Ile Trp Asp Asn Glu Thr Pro Ile Cys Asp Arg Ile Pro Cys Gly Leu
264 115 120 125
266 Pro Pro Thr Ile Thr Asn Gly Asp Phe Ile Ser Thr Asn Arg Glu Asn
267 130 135 140
269 Phe His Tyr Gly Ser Val Val Thr Tyr Arg Cys Asn Pro Gly Ser Gly
270 145 150 155 160
272 Gly Arg Lys Val Phe Glu Leu Val Gly Glu Pro Ser Ile Tyr Cys Thr

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/612,314

DATE: 11/14/2001

TIME: 09:58:50

Input Set : A:\379450004.app

Output Set: N:\CRF3\11142001\I612314.raw

```

273          165          170          175
275 Ser Asn Asp Asp Gln Val Gly Ile Trp Ser Gly Pro Ala Pro Gln Cys
276          180          185          190
278 Ile Ile Pro Asn Lys Cys Cys Asp Gly Pro Lys Lys Lys Lys Lys
279          195          200          205
281 Ser Pro Ser Lys Ser Ser Gly
282      210          215
285 <210> SEQ ID NO: 9
286 <211> LENGTH: 231
287 <212> TYPE: PRT
288 <213> ORGANISM: Artificial Sequence
290 <220> FEATURE:
291 <221> NAME/KEY: UNSURE
292 <222> LOCATION: (214)..(215)
293 <223> OTHER INFORMATION: Residues 1-214 are a first polypeptide chain and
294 residues 215-231 are a second polypeptide chain
295 linked by a disulphide bond formed between the
296 cysteines at positions 214 and 215
298 <220> FEATURE:
299 <221> NAME/KEY: UNSURE
300 <222> LOCATION: (231)
301 <223> OTHER INFORMATION: The C-terminal glycine is linked to an
302 N-[Myristoyl] group
304 <220> FEATURE:
305 <223> OTHER INFORMATION: Description of Artificial Sequence: [SCRL-3/switch
306 fusion] disulphide linked to [MSWP-1]
308 <220> FEATURE:
309 <221> NAME/KEY: UNSURE
310 <222> LOCATION: (214)
311 <223> OTHER INFORMATION: The cysteine at position 214 is -Cys-COOH
313 <220> FEATURE:
314 <221> NAME/KEY: UNSURE
315 <222> LOCATION: (215)
316 <223> OTHER INFORMATION: The cysteine at position 215 is CONH2-Cys-
318 <400> SEQUENCE: 9
319 Met Gln Cys Asn Ala Pro Glu Trp Leu Pro Phe Ala Arg Pro Thr Asn
320      1          5          15
322 Leu Thr Asp Glu Phe Glu Phe Pro Ile Gly Thr Tyr Leu Asn Tyr Glu
323          20          25          30
325 Cys Arg Pro Gly Tyr Ser Gly Arg Pro Phe Ser Ile Ile Cys Leu Lys
326          35          40          45
328 Asn Ser Val Trp Thr Gly Ala Lys Asp Arg Cys Arg Arg Lys Ser Cys
329          50          55          60
331 Arg Asn Pro Pro Asp Pro Val Asn Gly Met Val His Val Ile Lys Gly
332      65          70          75          80
334 Ile Gln Phe Gly Ser Gln Ile Lys Tyr Ser Cys Thr Lys Gly Tyr Arg
335          85          90          95
337 Leu Ile Gly Ser Ser Ser Ala Thr Cys Ile Ile Ser Gly Asp Thr Val
338      100          105          110

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/612,314

DATE: 11/14/2001

TIME: 09:58:51

Input Set : A:\379450004.app

Output Set: N:\CRF3\11142001\I612314.raw